5.6 The Implementation of the 65nm Dual-Core 64b Merom Processor

Nabeel Sakran, Marcelo Yuffe, Moty Mehalel, Jack Doweck, Ernest Knoll, Avi Kovacs

Intel, Haifa, Israel

The Merom design implements the Core™ microarchitecture, which is the new foundation for a range of desktop, mobile, and mainstream server multi-core processors. Merom is designed for efficiency and optimized for performance across a range of market segments and power envelopes. The processor operates at a wide range of frequencies (1 to 3GHZ) and voltages (0.85 to 1.325V). Figure 5.6.1 shows the frequency, voltage and power envelopes of the supported products.

The processor consists of two single-threaded 64b cores and a 4MB shared L2 cache. Each core has wide dynamic execution including 4-wide decode/rename, 32-wide scheduler, deep out-of-order storage, as well as micro- and macro-ops fusion. An enhanced multi-media function which executes a 128b packed instruction in a single cycle, improving the throughput by 2×. An advanced memory and cache access, which eliminates false memory dependencies and provides high, sustained bandwidth of cache line (64B) every 2 cycles. Merom implements an enhanced power capability, which effectively manages the idle power consumption to maximize the battery life of mobile products (4 to 5 hours). The processor delivers 40% more performance while requiring 40% less power than the Pentium D processor. Figure 5.6.2 shows the block diagram of this processor. Merom is implemented in a 65nm CMOS process with 8 copper interconnect layers and low-k carbon-doped oxide (k=2.9) inter-level dielectric. The 143mm^2 die has 291M transistors.

Merom has a 16-way 4MB shared L2 cache. The cache access time is optimized to achieve 2.0ns delay from address-in to data-out. The access path includes tag lookup, data read and fully transparent error correction in both tag and data arrays. To reduce the cache leakage power, sleep transistors (STs) are implemented in memory arrays, decoders and write drivers. The ST is based on a PMOS device, creating a virtual V_{cc} level, which can get as low as 500mV below the chip's V_{cc} level while still retaining the data. This reduces the memory array leakage by 3×. Figure 5.6.3 shows the L2 block diagram.

Another use of the memory array ST is to provide a dynamic cache-by-demand scheme. In this scheme, the microarchitecture identifies low usage of the cache, and reduces the effective cache size by disabling parts of the cache. The STs of the disabled parts move from sleep-mode to shut-off-mode. Fully floating the virtual V_{cc} in shut-off mode reduces the array leakage by 7×.

The cache has a two-dimensional redundancy scheme that includes both row and column redundancy. The redundancy maximizes the product yield and enables operation at aggressive minimum voltage levels. The cache has a fully transparent error-correction mechanism, as well as a new error correction reporting algorithm. This enables reliable and full-performance operation of the cache even if a small number of cells are bad. To ensure that the cache operates within expected reliability targets, hardware counts the lines that incur repeated corrections, and raises a cautionary status indicator if this count exceeds a preset threshold. The error correction mechanism and the enhanced redundancy scheme help meet the 0.85V operating voltage.

Figure 5.6.4 shows the clock distribution map. Two cascaded PLLs generate the clock for the front-side bus (FSB), the two cores and the cache system. The first PLL is fed by a 133 to 333MHz external reference clock and drives the core PLL reference clock, which multiplies the frequency by the bus fraction and generates the core clock. The PLLs are powered by on-die voltage regulators that enable a power-supply rejection ratio better than 20dB in the bandwidth of DC to 200MHz. The clock distribution implements an active de-skewing scheme, using digital delay lines to minimize the skew across the die to less than 18ps.

The FSB is logic-compatible to the previous generation mobile, desktop and server FSB and supports data rates of 666MT/s in the mobile version to 1066MT/s in the desktop version and up to 1333MT/s in the server version.

Two Merom processors can be assembled together in a multi-chip module (MCM) to create a quad-core processor. The FSB of the two chips are connected at the package level such that the dual-core and the quad-core versions have the same footprint and pinout. The quad-core FSB frequency is the same as the dual-core version.

To allow smooth operation of the FSB I/O buffers in the different system environments, the processor has a buffer configuration mechanism that allows presetting of the electrical characteristics of the I/O buffers during the production-testing phase. Extensive system level simulations and post-silicon characterization is done to define the optimum operation point for every product flavor. Then an internal PROM memory is programmed during the testing phase. During the power up process, the CPU reads this PROM and calibrates the size of the output buffer driver, the pre-driver, the bus delay-line length and selects between three predefined on-die termination configurations. This scheme ensures that every product flavor is optimized to its platforms electrical environment.

The processor is assembled in a variety of packages that support the requirements of the different market segments: 478-pin PGA and BGA for mobile and 780-pin LGA for desktop and server, with or without an integrated heat spreader. The different flavors use between 198 and 225 signal pins, the rest are dedicated to power and ground.

The chip is powered by three different power supplies, a variable voltage power supply for the core and cache that supports the Speed Step™ technology, a 1.5V power supply for the analog circuits and a 1.05 to 1.2V power supply for the FSB circuitry.

Due to the high execution parallelism, it is difficult to predict the location of the thermal hot spots on the chip. To resolve this problem, the processor has four independent digital thermal sensors that are independently calibrated at the production testing phase. The internal logic of the chip polls the thermal sensors and reports the temperature of the hottest spot to the thermal management software. An analog thermal probe is also supported. Figure 5.6.5 shows an example of a measured thermal map and the relative location of the thermal sensors.

Merom implements several techniques for speed debug and test shown in Fig. 5.6.6. The locate critical paths (LCP) mode, described in Fig. 5.6.6a, locates the source and destination clock drivers that cause a speed-limiting path. The LCP concept is based on a delay activated in the local clock drivers and is controlled by a configuration register programmed by the TAP. The on-die functional redundancy check (FRC) mode, described in Fig. 5.6.6b, runs a single-thread test on both cores in parallel in a lock-step mode, with both cores behaving identically on a cycle-by-cycle basis. This mode reduces the test cost by half. Figure 5.6.7 is the die photo with the major blocks highlighted.

Acknowledgements:
The authors gratefully acknowledge the work of the talented and dedicated Intel team that implemented this processor.

References:
[1] P. Bai, C. Auth, S. Balakrishnan, et al., "A 65nm Logic Technology Featuring 35nm Gate Lengths, Enhanced Channel Strain, 8 Cu Interconnect Layers, Low-k ILD and 0.57μm^2 SRAM Cell," *IEDM Tech. Dig*., pp. 657-660, Dec., 2004.
[2] S. Rusu, S. Tam, H. Muljono, et al., "A Dual-Core Multi-Threaded Xeon Processor with 16MB L3 Cache," *ISSCC Tech. Digest*, pp. 118-119, Feb., 2006.

Attribute \ Market	Servers	Desktops	Mobility
Max frequency	3.00 GHz	2.93 GHz	2.33 GHz
Max Front Side Bus frequency	1.333 GT/s	1.066 GT/s	0.667 GT/s
Thermal Design Power	80W @ 3GHz 65 W @ lower frequencies	75W @ 2.93GHz 65W @ lower frequencies	34 W
Max VID	1.325V	1.325 V	1.200 V

Figure 5.6.1: Product Specifications.

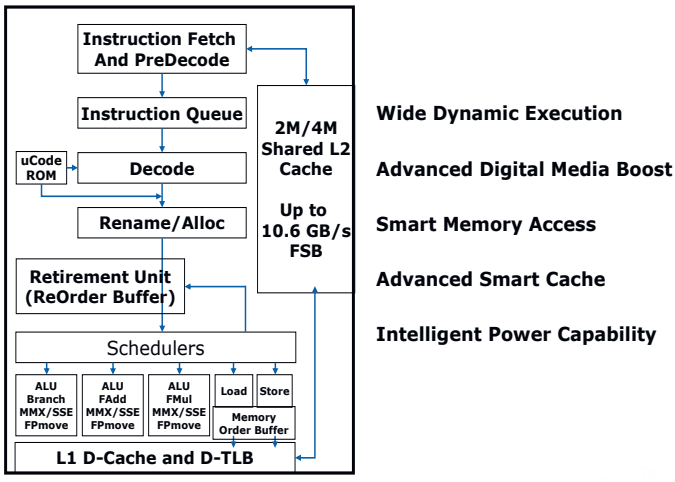

Figure 5.6.2: Core™ Microarchitecture main functions.

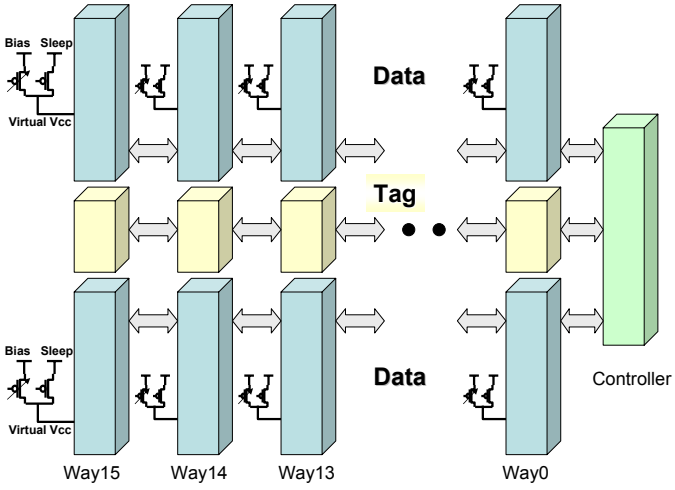

Figure 5.6.3: 4MB Shared L2 Block-Diagram.

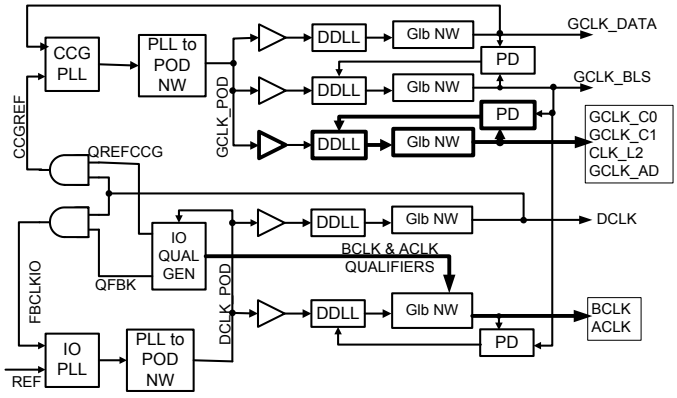

Figure 5.6.4: Clock Generation And Distribution.

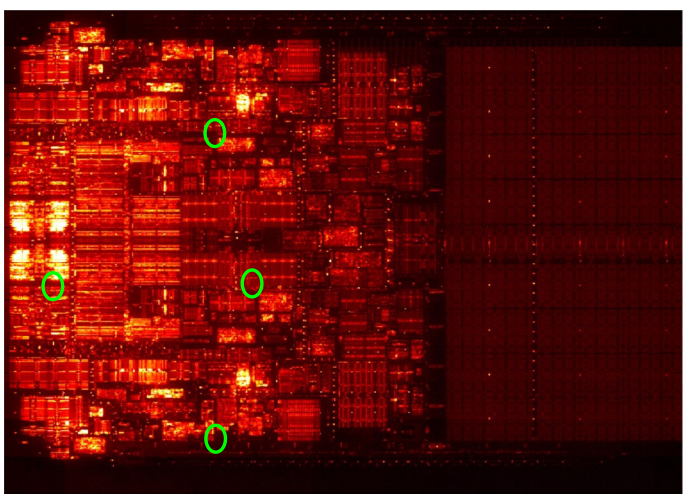

Figure 5.6.5: Merom thermal map.

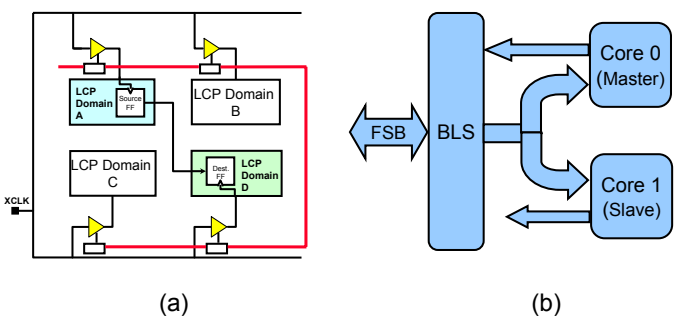

Figure 5.6.6: Speed debug / test techniques.

Continued on Page 590

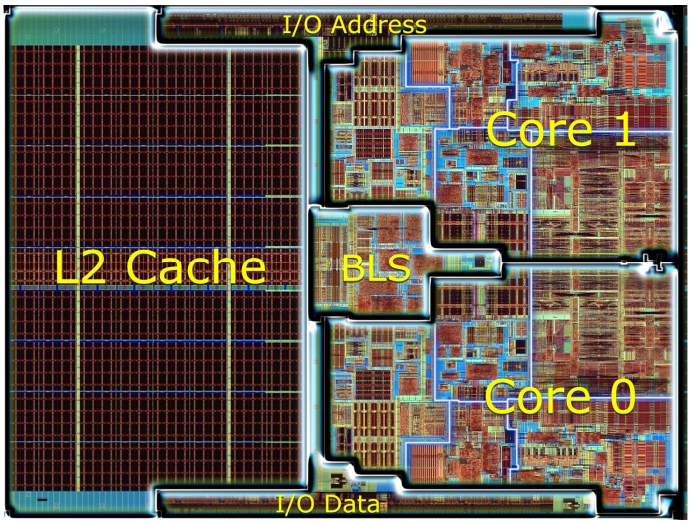

Figure 5.6.7: Die photo.

